1647

```
           RAW SEQUENCE LISTING              DATE: 07/06/2000
           PATENT APPLICATION: US/09/431,843B    TIME: 13:04:59

           Input Set : A:\13038.asc
           Output Set: N:\CRF3\07062000\I431843B.raw

 4  <110> APPLICANT: Zagon S., Ian
 5         Verderame, Michael
 6         Allen, Sandra
 7         McLaughlin J., Patricia
 9  <120> TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES ENCODING OPIOID GROWTH
10         FACTOR RECEPTORS
12  <130> FILE REFERENCE: Penn State
14  <140> CURRENT APPLICATION NUMBER: 09/431,843B
15  <141> CURRENT FILING DATE: 1999-11-02
17  <160> NUMBER OF SEQ ID NOS: 18
19  <170> SOFTWARE: PatentIn Ver. 2.1
21  <210> SEQ ID NO: 1
22  <211> LENGTH: 2250
23  <212> TYPE: DNA
24  <213> ORGANISM: Rattus norvegicus
26  <400> SEQUENCE: 1
27  tgggctcagc cacgccccag ggtgcccca gtgggactag ttcttcattc tggcagctgc  60
29  acacatctgt cagtgaggga atgtcaggtc tctcactctc ctctcctcac tatcctttcc  120
31  gcagaaagcg ggtcctcctg cttgtcgagt atggacgacc cggactgcga ttccacctgg  180
33  gaggaggaga gcgaggagga tggcgaggat ggccaggcgg atgatacgac cgatgaggac  240
35  acgggcgacg atgacggcga cgcggaggag gcacggccaa gcctgttcca gtccaggatg  300
37  acagggtacc gaaactggcg tgctatgcag gacatgcaaa gataccggca caactacccg  360
39  gatttgacag atcaagactg caatggtgac atgtgcaacc tgagcttcta caaaaatgag  420
41  atctgcttcc agccaaatgg ggctctcatc gaggacattc ttcagaactg gaaagacaac  480
43  tatgacctcc tggaagagaa tcactcctac atccagtggc tgtttcctct gcgggaacca  540
45  ggagtgaact ggcacgccaa gcccctcacc ctgaaggagg ttgaggcatt taaaagctcc  600
47  aaggaagtca gagagcgtct tgtccgggcc tatgagctca tgctgggctt ctatgggttc  660
49  caccttgagg accgggggcac gggtgctgta tgccgtgcac agaacttcca gccgcgcttc  720
51  cacaatctga acagccacag ccacaacaac ctgcgtatta cacgcatcct caagtcactg  780
53  ggtgagctgg gcttagaaca ctaccaggca cccctggtcc gcttcttcct ggaggagacc  840
55  cttgtacagc acaaactgcc cagcgtgcgc cagagtgccc tggactactt cctgttcgct  900
57  gtgcgctgcc ggcaccagcg ccgggagctt gtgtactttg cctgggagca cttcaagcct  960
59  cgccgagagt ttgtctgggg gccccgtgac aagctgaagc ccagaccata  1020
61  ccccagccac tgacgggacc agggcaggca gataaagatg agggctccag ggaccccctcc  1080
63  caagaggctg gcacccaggg tcggacctgt ggatctggaa gggacctgag tggggacagt  1140
65  ggaacagctg aggatccctc actgctgaac acaaagccct cagatggggg aaccttggat  1200
67  gggaaccaga gggatgaagc taagtccctg agtcccaagg agagcaagaa aaggaagttg  1260
69  gaggggaaca ggcaggagca ggtcccaggg gaggcagatc cccaggg tgt ctctgaggta  1320
71  gagaaaattg cccttaacct tgaggagtgt gcccttagcc ctatcagcca ggagcccagg  1380
73  gaggctgaac cgccctgtcc tgtggccagg gtggctaatg aggtaagaaa gcggaggaag  1440
75  gtggaggaag gggctgaggg tgatggagta gtcagtaaca ctcaaatgca ggccagtgcc  1500
77  ctgcctccta cccccttcaga gtgtcctgag gcccaaaagg atgggaatgg gccagaggac  1560
79  tcaaacagcc aggttggggc agaggattcc aaaagccagg tggggccgga ggatccaaac  1620
81  agccaggtgg ggctggagga cccaaacagc caggtcgggc cagaggaccc aaacagccag  1680
83  gtcgggccag aggacccaaa cagccaggtc gggccagagg acccaaacag ccaggtcggg  1740
85  ccagaggacc caaacagcca ggtggtgggg ccagagcaag ctgcctctaa gagccctgtg  1800
87  gaggaccctg actctgacac tatgggaacc tcagtggatg agtcagagga gttggcaagg  1860
```

RAW SEQUENCE LISTING   DATE: 07/06/2000
PATENT APPLICATION: US/09/431,843B   TIME: 13:04:59

Input Set : A:\13038.asc
Output Set: N:\CRF3\07062000\I431843B.raw

```
 89 attgaggcct ctgctgaacc cccaaagcct tagaggtgca tctcagtcct actcagccca 1920
 91 ctgcaggggg tttctgagtc cagagctctg ccgtaggctc ttcttggtgc cccacagtgc 1980
 93 tggcctctcc ctagtggtca ctgaggtggc caccagaggg actgaggccc tgccctcagg 2040
 95 gaaggccaag gccttcagaa ccctccttac ctcactgtgt cctcctccac tgccctctga 2100
 97 gccctgcgtt gtgatcagac cctaagggtc tagagggagg ggcctcttca ttagtctggt 2160
 99 gccaagtgag gccttttctg aataaactct ttagactttg tcaaaaaaaa aaaaaaaaaa 2220
101 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa                                 2250
105 <210> SEQ ID NO: 2
106 <211> LENGTH: 580
107 <212> TYPE: PRT
108 <213> ORGANISM: Rattus norvegicus
110 <400> SEQUENCE: 2
111 Met Asp Asp Pro Asp Cys Asp Ser Thr Trp Glu Glu Glu Ser Glu Glu
112   1               5                  10                  15
114 Asp Gly Glu Asp Gly Gln Ala Asp Asp Thr Thr Asp Glu Asp Thr Gly
115              20                  25                  30
117 Asp Asp Asp Gly Asp Ala Glu Glu Ala Arg Pro Ser Leu Phe Gln Ser
118          35                  40                  45
120 Arg Met Thr Gly Tyr Arg Asn Trp Arg Ala Met Gln Asp Met Gln Arg
121      50                  55                  60
123 Tyr Arg His Asn Tyr Pro Asp Leu Thr Asp Gln Asp Cys Asn Gly Asp
124  65                  70                  75                  80
126 Met Cys Asn Leu Ser Phe Tyr Lys Asn Glu Ile Cys Phe Gln Pro Asn
127              85                  90                  95
129 Gly Ala Leu Ile Glu Asp Ile Leu Gln Asn Trp Lys Asp Asn Tyr Asp
130              100                 105                 110
132 Leu Leu Glu Glu Asn His Ser Tyr Ile Gln Trp Leu Phe Pro Leu Arg
133          115                 120                 125
135 Glu Pro Gly Val Asn Trp His Ala Lys Pro Leu Thr Leu Lys Glu Val
136      130                 135                 140
138 Glu Ala Phe Lys Ser Ser Lys Glu Val Arg Glu Arg Leu Val Arg Ala
139 145                 150                 155                 160
141 Tyr Glu Leu Met Leu Gly Phe Tyr Gly Phe His Leu Glu Asp Arg Gly
142              165                 170                 175
144 Thr Gly Ala Val Cys Arg Ala Gln Asn Phe Gln Pro Arg Phe His Asn
145              180                 185                 190
147 Leu Asn Ser His Ser His Asn Asn Leu Arg Ile Thr Arg Ile Leu Lys
148          195                 200                 205
150 Ser Leu Gly Glu Leu Gly Leu Glu His Tyr Gln Ala Pro Leu Val Arg
151      210                 215                 220
153 Phe Phe Leu Glu Glu Thr Leu Val Gln His Lys Leu Pro Ser Val Arg
154 225                 230                 235                 240
156 Gln Ser Ala Leu Asp Tyr Phe Leu Phe Ala Val Arg Cys Arg His Gln
157              245                 250                 255
159 Arg Arg Glu Leu Val Tyr Phe Ala Trp Glu His Phe Lys Pro Arg Arg
160              260                 265                 270
162 Glu Phe Val Trp Gly Pro Arg Asp Lys Leu Arg Arg Phe Lys Pro Gln
163          275                 280                 285
165 Thr Ile Pro Gln Pro Leu Thr Gly Pro Gly Gln Ala Asp Lys Asp Glu
```

```
166      290                    295                    300
168 Gly Ser Arg Asp Pro Ser Gln Glu Ala Gly Thr Gln Gly Arg Thr Cys
169 305                    310                    315                    320
171 Gly Ser Gly Arg Asp Leu Ser Gly Asp Ser Gly Thr Ala Glu Asp Pro
172              325                    330                    335
174 Ser Leu Leu Asn Thr Lys Pro Ser Asp Gly Gly Thr Leu Asp Gly Asn
175              340                    345                    350
177 Gln Arg Asp Glu Ala Lys Ser Leu Ser Pro Lys Glu Ser Lys Lys Arg
178         355                    360                    365
180 Lys Leu Glu Gly Asn Arg Gln Glu Gln Val Pro Gly Glu Ala Asp Pro
181      370                    375                    380
183 Gln Gly Val Ser Glu Val Glu Lys Ile Ala Leu Asn Leu Glu Glu Cys
184 385                    390                    395                    400
186 Ala Leu Ser Pro Ile Ser Gln Glu Pro Arg Glu Ala Glu Pro Pro Cys
187              405                    410                    415
189 Pro Val Ala Arg Val Ala Asn Glu Val Arg Lys Arg Arg Lys Val Glu
190              420                    425                    430
192 Glu Gly Ala Glu Gly Asp Gly Val Val Ser Asn Thr Gln Met Gln Ala
193         435                    440                    445
195 Ser Ala Leu Pro Pro Thr Pro Ser Glu Cys Pro Glu Ala Gln Lys Asp
196      450                    455                    460
198 Gly Asn Gly Pro Glu Asp Ser Asn Ser Gln Val Gly Ala Glu Asp Ser
199 465                    470                    475                    480
201 Lys Ser Gln Val Gly Pro Glu Asp Pro Asn Ser Gln Val Gly Leu Glu
202              485                    490                    495
204 Asp Pro Asn Ser Gln Val Gly Pro Glu Asp Pro Asn Ser Gln Val Gly
205              500                    505                    510
207 Pro Glu Asp Pro Asn Ser Gln Val Gly Pro Glu Asp Pro Asn Ser Gln
208         515                    520                    525
210 Val Gly Pro Glu Asp Pro Asn Ser Gln Val Val Gly Pro Glu Gln Ala
211      530                    535                    540
213 Ala Ser Lys Ser Pro Val Glu Asp Pro Asp Ser Asp Thr Met Gly Thr
214 545                    550                    555                    560
216 Ser Val Asp Glu Ser Glu Glu Leu Ala Arg Ile Glu Ala Ser Ala Glu
217              565                    570                    575
219 Pro Pro Lys Pro
220              580
223 <210> SEQ ID NO: 3
224 <211> LENGTH: 987
225 <212> TYPE: DNA
226 <213> ORGANISM: Rattus norvegicus
228 <220> FEATURE:
229 <221> NAME/KEY: unsure
230 <222> LOCATION: (164)
231 <223> OTHER INFORMATION: n is unsure
233 <220> FEATURE:
234 <221> NAME/KEY: unsure
235 <222> LOCATION: (179)
236 <223> OTHER INFORMATION: n is unsure
```

```
           RAW SEQUENCE LISTING            DATE: 07/06/2000
           PATENT APPLICATION:  US/09/431,843B     TIME: 13:04:59

           Input Set : A:\13038.asc
           Output Set: N:\CRF3\07062000\I431843B.raw

       238 <220> FEATURE:
       239 <221> NAME/KEY: unsure
       240 <222> LOCATION: (184)
       241 <223> OTHER INFORMATION: n is unsure
       243 <220> FEATURE:
       244 <221> NAME/KEY: unsure
       245 <222> LOCATION: (213)
       246 <223> OTHER INFORMATION: n is unsure
       248 <220> FEATURE:
       249 <221> NAME/KEY: unsure
       250 <222> LOCATION: (240)
       251 <223> OTHER INFORMATION: n is unsure
       253 <220> FEATURE:
       254 <221> NAME/KEY: unsure
       255 <222> LOCATION: (555)
       256 <223> OTHER INFORMATION: n is unsure
       258 <220> FEATURE:
       259 <221> NAME/KEY: unsure
       260 <222> LOCATION: (622)
       261 <223> OTHER INFORMATION: n is unsure
       263 <400> SEQUENCE: 3
       264 cattgggccg acgtcgcatg ctcctctaga ctcgaggaat tcgggcccca gggtgtctct 60
       266 gaggtagaga aaattgccct taaccttgag gagtgtgccc ttagccctat cagccaggag 120
W--> 268 cccagggagg stgaaccgcc ctgtcctgtg gccagggtgg ctanaatgag gtaagaaang 180
W--> 270 cggnaggaag gtggaggaag gggctgaggg tgnatggagt agtcagtaac actyaaatgn 240
       272 caggccagtg ccctgcctcc tacccttca gagtgtcctg aggcccaaaa ggatgggaat 300
       274 gggccagagg actcaaacag ccaggttggg gcagagatt ccaaaagcca ggtggggccg 360
       276 gaggatccaa acagccaggt ggggctggag gacccaaaca gccaggtcgg gccagaggac 420
       278 ccaaacagcc aggtcgggcc agaggaccca aacagccagg tcgggccaga ggacccaaac 480
       280 agccaggtcg ggccagagga cccaaacagc caggtggtgg ggcagagca agctgcctct 540
W--> 282 aagagccctg tgganggacc ctgactctga cactatggga acctcagtgg atgagtcaga 600
W--> 284 ggagttggca aggattgagg cntytgctga acccccaaag ccttagaggt gcatttcagt 660
       286 cctactcagc ccactgcagg gggtttctga gtccagagct ctgccgtagg ctcttcttgg 720
       288 tgccccacag tgctggcctc tccctastgg tcactgaggt ggccaccaga gggactgagg 780
       290 ccctgccctc agggaaggcc aaggccttca gaaccctcct tacctcactg tgtcctcctc 840
       292 cactgccctc tgagccctgc gttgtgatca gaccctaagg gtctagaggg aggggcctct 900
       294 tcattagtct ggtgccaagt gaggcctttt ctgaataaac tctttagact ttgtcaaaaa 960
       296 aaaaaaaaaa aaaaaaaaaa aaaaaaa                                      987
       300 <210> SEQ ID NO: 4
       301 <211> LENGTH: 2290
       302 <212> TYPE: DNA
       303 <213> ORGANISM: Homo sapiens
       305 <400> SEQUENCE: 4
       306 tagaattcag cggccgctga attctagccg agcatggacg accccgactg cgactccacc 60
       308 tgggaggagg acgaggagga tgcggaggac gcggaggacg aggactgcga ggacggcgag 120
       310 gccgccggcg cgagggacgc ggacgcaggg gacgaggacg aggagtcgga gggagccgcgg 180
       312 gcggcgcggc ccagctcgtt ccagtccaga atgacagggt ccagaaactg gcgagccacg 240
       314 agggacatgt gtaggtatcg gcacaactat ccggatctgg tggaacgaga ctgcaatggg 300
       316 gacacgccaa acctgagttt ctacagaaat gagatccgct tcctgcccaa cggctgtttc 360
```

```
RAW SEQUENCE LISTING                    DATE: 07/06/2000
PATENT APPLICATION:  US/09/431,843B     TIME: 13:04:59

Input Set : A:\13038.asc
Output Set: N:\CRF3\07062000\I431843B.raw

318 attgaggaca ttcttcagaa ctggacggac aactatgacc tccttgagga caatcactcc 420
320 tacatccagt ggctgtttcc tctgcgagaa ccaggagtga actggcatgc caagcccctc 480
322 acgctcaggg aggtcgaggt gtttaaaagc tcccaggaga tccaggagcg gcttgtccgg 540
324 gcctacgagc tcatgctggg cttctacggg atccggctgg aggaccgagg cacgggcacg 600
326 gtgggccgag cacagaacta ccagaagcgc ttccagaacc tgaactggcg cagccacaac 660
328 aacctccgca tcacacgcat cctcaagtcg ccgtgtgagc tgagcctcga gcacttccag 720
330 gcgccactgg tccgcttctt cctggaggag acgctggtgc ggcgggagct gccgggggtg 780
332 cggcagagtg ccctggacta cttcatgttc gccgtgcgct gccgacacca cgccgccag 840
334 ctggtgcact tcgcctggga gcacttccgg ccccgctgca agttcgtctg ggggcccca 900
336 gacaagctgc ggaggttcaa gcccagctct ctgccccatc cgctcgaggg ctccaggaag 960
338 gtggaggagg aaggaagccc cggggacccc gaccacgagg ccagcaccca gggtcggacc 1020
340 tgtgggccag agcatagcaa gggtggggc agggtggacg aggggcccca gccacggagc 1080
342 gtggagcccc aggatgcggg acccctggag aggagccagg gggatgaggc aggggggccac 1140
344 gggaagata ggccggagcc cttaagcccc aaagagagca agaagaggaa gctggagctg 1200
346 agccggcggg agcagccgcc cacagagcca ggccctcaga gtgcctcaga ggtggagaag 1260
348 atcgctctga atttggaggg gtgtgccctc agccagggca gcctcaggac ggggacccag 1320
350 gaagtgggcg gtcaggaccc tggggaggca gtgcagccct gccgccaacc cctgggagcc 1380
352 agggtggccg acaaggtgag gaagcggagg aaggtggatg agggtgctgg ggacagtgct 1440
354 gcggtggcca gtggtggtgc ccagaccttg gcccttgccg ggtcccctgc cccatcgggg 1500
356 caccccaagg ctggacacag tgagaacggg gttgaggagg acacagaagg tcgaacgggg 1560
358 cccaaagaag gtacccctgg gagcccatcg gagacccag gccccagccc agcaggacct 1620
360 gcaggggacg agccagccga gagcccatcg gagaccccag gcccccgccc ggcaggacct 1680
362 gcaggggacg agccagccga gagcccatcg gagacccag gcccccagccc ggcaggacct 1740
364 acaagggatg agccagccga gagcccatcg gagaccccag gcccccgccc ggcaggacct 1800
366 gcaggggacg agccagccga gagcccatcg gagacccag gcccccgccc ggcaggacct 1860
368 gcaggggacg agccagccga gagcccatcg gagacccag gcccccagccc ggcaggacct 1920
370 acaagggatg agccagccaa ggcggggag gcagcagagt tgcagkacgc agaggtggag 1980
372 tcttctgcca agtctgggaa gccttaagga aaggagtgcc cgtcggcgtc ttggtcctcc 2040
374 tgtccctgct gcaggggctg gggcctccgg agcttgctgc gggctcccct caggctctgc 2100
376 ttcgtgaccc gtgacccatg acccacagtg ctggcctcct gtggggccac tatagcarsc 2160
378 accagaagcc gcgaggccct cagggaagcc caaggcctgc agaagcctcc tggcctggct 2220
380 gtgtcttccc cacccagctc tcccctgcgc ccctgtcttt gtaaattgac ccttctggag 2280
382 tgggggggcgg                                                      2290
386 <210> SEQ ID NO: 5
387 <211> LENGTH: 2408
388 <212> TYPE: DNA
389 <213> ORGANISM: Homo sapiens
391 <400> SEQUENCE: 5
392 tagaattcag cggccgctga attctagccg agcatggacg accccgactg cgactccacc 60
394 tgggaggagg acgaggagga tgcggaggac gcggaggacg aggactgcga ggacggcgag 120
396 gccgccggcg cgagggacgc ggacgcaggg gacgagtcgga aggagtcgga ggacggcggg 180
398 gcggcgcggc ccagctcgtt ccagtccaga atgacagggt ccagaaactg gcgagccacg 240
400 agggacatgt gtaggtatcg gcacaactat ccggatctgg tggaacgaga ctgcaatggg 300
402 gacacgccaa acctgagttt ctacagaaat gagatccgct tcctgcccaa cggctgtttc 360
404 attgaggaca ttcttcagaa ctggacggac aactatgacc tccttgagga caatcactcc 420
406 tacatccagt ggctgtttcc tctgcgagaa ccaggagtga actggcatgc caagcccctc 480
408 acgctcaggg aggtcgaggt gtttaaaagc tcccaggaga tccaggagcg gcttgtccgg 540
410 gcctacgagc tcatgctggg cttctacggg atccggctgg aggaccgagg cacgggcacg 600
412 gtgggccgag cacagaacta ccagaagcgc ttccagaacc tgaactggcg cagccacaac 660
```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY   DATE: 07/06/2000
PATENT APPLICATION: US/09/431,843B   TIME: 13:05:00

Input Set : A:\13038.asc
Output Set: N:\CRF3\07062000\I431843B.raw

L:268 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:270 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:282 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:284 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:1227 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13